A robust stochastic filter for point tracking in image sequences

Élise Arnaud, Etienne Mémin, Bruno Cernuschi-Frias

A ROBUST STOCHASTIC FILTER FOR POINT TRACKING IN IMAGE SEQUENCES

E. Arnaud, E. Mémin

IRISA, Université de Rennes 1
Campus de Beaulieu, 35 042 Rennes, France
{earnaud,memin}@irisa.fr

B. Cernuschi-Frías

LIPSIRN, Universidad de Buenos Aires
Argentina

ABSTRACT

The approach we investigate for point tracking combines within a stochastic filtering framework a dynamic model relying on the optical flow constraint and measurements provided by a matching technique. Focusing on points belonging to regions described by a global dominant motion, the proposed tracking system is linear. Since we focus on the case where the system depends on the images, the tracker is built from a Conditional Linear Filter, derived through the use of a conditional linear minimum variance estimator. This conditional tracker authorizes to significantly improve results in some general situation. In particular, such an approach allows us to deal in a simple way with the tracking of points following trajectories with abrupt changes and occlusions.

1. INTRODUCTION

The objective of point tracking in computer vision consists in reconstructing a point trajectory along a given image sequence. This problem is a basic but essential process, on which many high level tasks depend, such as motion estimation, registration, 3D reconstruction, etc.

To track a point, it is necessary to make some conservation assumptions on some point related information. These hypotheses may concern the point motion, or a photometric invariance in a neighborhood of the point. Nevertheless, any dynamic prior model on the point motion is very difficult to establish without any *a priori* knowledge on the dynamic of the surrounding object. These intrinsic difficulties have driven researchers to implement local techniques based on geometric and luminance invariants which locally characterize the gray value signal. The assumption of luminance pattern conservation along a trajectory has led to devise two kinds of methods. The first ones are intuitive methods based on correlation. Although the used similarity functions are not invariant to affine image transformations, these methods remain very popular [1]. The second ones are defined as differential trackers, built on a differential formulation of a similarity criterion. In particular, the well-known Shi-Tomasi-Kanade (STK) tracker [2] is constructed from the sum-of-squared-differences criterion.

In order to properly mix these two sources of information, we settle their competition into a stochastic filtering modelization. Such a framework models the problem by a discrete hidden Markov chain, described by a system. This system consists of a state equation, which characterizes the evolution law of the state to be estimated, and a measure equation which links the observation to the state. Stochastic filters give thus procedures to estimate the distribution probability of the state conditionally to all past measures. These filters, such as Kalman filter in the linear Gaussian case [3, 4] or sequential Monte Carlo approximation methods in the nonlinear case [5] are well-known to improve tracker robustness to outliers and occlusions.

In our case, the whole system which describes the point tracking problem depends on the image sequence. Indeed, both dynamic and measures are extracted from the image sequence at each discrete instant. They rely on the one hand on a differential method and on the other hand on a correlation criterion. One key-point of our work is therefore to propose a well-founded filtering framework allowing to deal with such a usual situation. The resulting filters are built following the traditional setup of stochastic filters, by considering a conditioning with respect to the image sequence data. Such a conditioning obliges us to adapt the usual estimators used in tracking applications. In this paper, we focus on the linear version of these filters, named *Conditional Linear Filter* (CLF). It is important to note that the CLF allows to simply deal with linear *a priori*-free models. This filter is described section 2.

The use of the Conditional Linear Filter for point tracking is described section 3. The presented system is particularly well-suited to image sequences exhibiting dominant motion situations. Indeed, it enables us to combine global and local pieces of information. Moreover, an automatic computation of the measure noise covariance leads the tracker to be robust to occlusions, image noise and abrupt changes of trajectories. Section 4 presents results on real-world sequences, proving the ability of the CLF to deal with difficult situations. The proposed filter is compared to the Shi-Tomasi-Kanade tracker [2] and to a nonlinear particle filter tracker.

2. CONDITIONAL LINEAR FILTER

For the sake of clarity, let us first define the notations used throughout this paper. Let $\mathbf{I}_k$ denote a random variable which corresponds to an image obtained at time k. The finite sequence of variables $\{\mathbf{I}_k, k = 0, ..., n\}$ will be represented by $\mathbf{I}_{0:n}$. The discrete hidden Markov state process is denoted $\mathbf{x}_{0:n} = \{\mathbf{x}_0, \mathbf{x}_1, ..., \mathbf{x}_n\}$ and the set of observations $\mathbf{z}_{1:n} = \{\mathbf{z}_1, \mathbf{z}_1, ..., \mathbf{z}_n\}$. At each time, the classical filtering problem consists in having an accurate approximation of the posterior probability density of state $\mathbf{x}_k$ given the whole set of past and present measures $\mathbf{z}_{1:k}$.

As introduced in section 1, in our point of view, tracking features from image sequences may require in some cases to define a slightly modified framework of stochastic filtering. A first problem comes from the choice of the observation model. As a matter of fact, the measure on which one should ideally rely is the image sequence itself. Unfortunately, images have too large dimensions and too complex structures to be used directly. Therefore, one usually defines a digest structured observation built *from* the images.

Another source of difficulties comes from the definition of appropriate dynamic models. These models are usually defined from *a priori* models [6, 7] or obtained by learning [5]. A severe limitation of these models arises facing the tracking of features whose trajectories exhibit abrupt changes and occlusions or simply obey too complex dynamic laws, which can not be learned or predicted. This is the case when tracking very general (punctual / picked) entities in images of any kinds. In such a context, the only possibility is to rely on a dynamical model extracted from the image sequence itself. As a result, we face a tracking problem for which the whole system (observation and state) depends on the images. In that peculiar case, we propose here a modified formulation of a linear filtering approach.

Knowing a realization of $\mathbf{I}_{0:k}$, the linear image based tracking system is modeled as follows :

$$\mathbf{x}_k = A_k^{\mathbf{I}_{0:k}} \mathbf{x}_{k-1} + \mathbf{b}_k^{\mathbf{I}_{0:k}} + \mathbf{w}_k^{\mathbf{I}_{0:k}} \tag{1}$$

$$\mathbf{z}_k = H_k^{\mathbf{I}_{0:k}} \mathbf{x}_k + \mathbf{v}_k^{\mathbf{I}_{0:k}} \tag{2}$$

The index $\mathbf{I}_{0:k}$ indicates a possible dependence on the image data. Let us remind that in our case, matrices $A_k^{\mathbf{I}_{0:k}}$, $H_k^{\mathbf{I}_{0:k}}$ and vector $\mathbf{b}_k^{\mathbf{I}_{0:k}}$ may be estimated from $\mathbf{I}_{0:k}$. The state noise $\mathbf{w}_k^{\mathbf{I}_{0:k}}$ and the measure noise $\mathbf{v}_k^{\mathbf{I}_{0:k}}$ may as well be specified from $\mathbf{I}_{0:k}$. They are supposed to be zero mean independent white noises, possibly non Gaussian, of known conditional covariances respectively denoted $Q_k^{\mathbf{I}_{0:k}}$ and $R_k^{\mathbf{I}_{0:k}}$. To be able to tackle the problem of the dependency on the sequence, the Conditional Linear Filter is derived through an extension of the linear minimum variance estimator. This estimator that we call the *conditional linear minimum variance estimator* provides us an estimation of the two first moments of $p(\mathbf{x}_k|\mathbf{z}_{1:k}, \mathbf{I}_{0:k})$. This description is obviously sufficient to have the entire knowledge of the expected density if the linear model is Gaussian. For non-Gaussian noises, this description provides only a Gaussian approximation of the posterior density function.

2.1. Conditional linear minimum variance estimator

As already said, the Conditional Linear Filter is built by relying on a conditional linear minimum variance estimator:

Definition 1 *Let X,Z,W be 3 jointly distributed random variables. $E_W^*[X|Z]$ denotes the best estimator of X, linear in Z, conditionally to W:*

$$E_W^*[X|Z] = AZ + B$$

with A and B such that $E[\|X - AZ - B\|^2|W]$ is minimum. $E_W^*[X|Z]$ is called the conditional linear minimum variance estimator.

It must be noticed that $E_W^*[X|Z]$ is not an expectation. Denoting $\Sigma_{X,Z|W} = E[XZ^t|W] - E[X|W]E[Z|W]^t$, the following important result is obtained (see appendix):

$$E_W^*[X|Z] = E[X|W] + \Sigma_{X,Z|W} \Sigma_{Z,Z|W}^{-1} (Z - E[Z|W]).$$

It can be checked that this estimator shares some similar properties to the linear minimum variance estimator.

2.2. Tracking with Conditional Linear Filter

We consider a linear model of the form described in (1,2). To simplify the notations, the index $\mathbf{I}_{0:k}$ will be omitted in the following of this section. Let us denote $\hat{\mathbf{x}}_{k+1|k} = E_{\mathbf{I}_{0:k+1}}^*[\mathbf{x}_{k+1}|\mathbf{z}_{1:k}]$ and $\Sigma_{k+1|k}$ the associated conditional error covariance. Considering conditional expressions induced by E^*, and relying on its properties, a recursive formulation of $\hat{\mathbf{x}}_{k+1|k}$ can be found through similar manipulations to the usual Gaussian case:

$$\hat{\mathbf{x}}_{k+1|k} = A_{k+1}\, \hat{\mathbf{x}}_{k|k-1} + \mathbf{b}_{k+1} + \widetilde{K}_k\, (\mathbf{z}_k - H_k\, \hat{\mathbf{x}}_{k|k-1}),$$

where matrix $\widetilde{K}_k$ is defined using the Kalman gain K_k:

$$\begin{aligned} \widetilde{K}_k &= A_{k+1} K_k \\ &= A_{k+1} (\Sigma_{k|k-1} H_k^t)(H_k \Sigma_{k|k-1} H_k^t + R_k)^{-1}. \end{aligned}$$

A recursive expression of the conditional estimation error covariance $\Sigma_{k+1|k}$ can also be obtained as:

$$\begin{aligned} \Sigma_{k+1|k} = {} & (A_{k+1} - \widetilde{K}_k\, H_k)\, \Sigma_{k|k-1}\, (A_{k+1} - \widetilde{K}_k\, H_k)^t \\ & + Q_{k+1} + \widetilde{K}_k\, R_k\, \widetilde{K}_k^t. \end{aligned}$$

These equations can be further split to distinguish the prediction step and the update step. In order to limit the computational cost, it may be useful to define a research area where the estimation process of the measure $\mathbf{z}_k$ is applied. Such a region, called validation gate, is defined as an area of the measure space where the future observation will be found with some high probability. Gates are generally used in radar tracking problems, for clutter reduction [8]. They are here defined through the use of the probability distribution $p(\mathbf{z}_k|\mathbf{z}_{1:k-1}, \mathbf{I}_{0:k})$, that is approximated by a normal distribution (this expression is exact in case of Gaussian noises). An ellipsoidal probability concentration region is then defined as:

$$gate_k = \{\mathbf{z}_k : \epsilon_k = \widetilde{\mathbf{z}}^t_{k|k-1} S_k^{-1} \widetilde{\mathbf{z}}_{k|k-1} \leq \gamma\},$$

where $\widetilde{\mathbf{z}}_{k|k-1} = \mathbf{z}_k - \hat{\mathbf{z}}_{k|k-1}$ is the innovation sequence ($\hat{\mathbf{z}}_{k|k-1} = H_k \hat{\mathbf{x}}_{k|k-1}$ is the predictive measure) and $S_k = H_k \Sigma_{k|k-1} H_k^t + R_k$ its associated conditional covariance.

The resulting filter constitutes a tracker which resembles the Kalman filter for Gaussian linear models. It is nevertheless important to note that (i) the use of Kalman recursive equations are now well justified for the specific case of image based systems through the use of a conditional minimum variance estimator and (ii) it povides a sound framework which enables a probabilistic competition between two estimation processes on the image sequence.

3. APPLICATION TO POINT TRACKING

Considering a given point in the scene, the problem of tracking this single feature in an image sequence can be defined as, at each time, estimating the position of the point projection in the image plane. We consider the most general context, where no knowledge on the dynamic of the surrounding object is available. As said before, the proposed solution to tackle the lack of *a priori* information is to compute the entire model from the image sequence and solve the system with the Conditional Linear Filter. In the considered tracking problem, each state $\mathbf{x}_k$ represents the location of the point projection at time k, in image $\mathbf{I}_k$, observable through the measure $\mathbf{z}_k$. Let us point out, that for the kind of system we focus on, both measures and the dynamic equation are built from $\mathbf{I}_{0:k}$. Indeed, we combine a dynamic model relying on the optical flow constraint and measures provided by a matching technique.

3.1. Conditional state equation

In order to be reactive to any change of speed and direction of the point, the conditional state equation defines the instantaneous model of the feature motion. It describes the motion of a point $\mathbf{x}_{k-1}$ between images $k-1$ and k, and gives an estimation of $\mathbf{x}_k$. The conditional state equation can be written as :

$$\mathbf{x}_k = \mathbf{x}_{k-1} + P(\mathbf{x}_{k-1})\, \boldsymbol{\theta}_k(\mathbf{x}_{k-1}) + \mathbf{w}_k$$

where $\mathbf{u}(\mathbf{s}) = P(\mathbf{s})\, \boldsymbol{\theta}_k(\mathbf{s})$ denotes the estimated motion vector of pixel $\mathbf{s} = (x, y)^t$, modeled as a 6-parameter affine motion. $\boldsymbol{\theta}_k(\mathbf{s})$ is the parameter vector, and $P(\mathbf{s})$ is a matrix related to the parametric model whose entries depend on spatial coordinates x and y.

A robust parametric motion estimation technique [9] allows to estimate reliably this parametric model representing the dominant apparent velocity field on a given support $\mathcal{R}$. The use of such a method on an appropriate support around $\mathbf{x}_{k-1}$, from images $\mathbf{I}_{k-1}$ and $\mathbf{I}_k$, provides an estimate of the motion vector $\mathbf{u}(\mathbf{x}_{k-1})$.

It is important to note that when the point motion corresponds to a global dominant motion, the estimation support is the whole image grid. Consequently, for a point $\mathbf{s}$ belonging to the dominant motion support, $\boldsymbol{\theta}_k(\mathbf{s})$ does not depend on $\mathbf{s}$, which induces the linearity of the state equation (let remark that it is not the case for points moving independently from the global motion). As a result, for these points, the obtained linear dynamic is of the form (1), where A_k is the matrix related to rotation, divergence and shear motion, $\mathbf{b}_k$ is a translation vector. A_k and $\mathbf{b}_k$ are estimated on the image sequence. The noise variable $\mathbf{w}_k$ accounts for errors related to the *global* motion and is likely to be non Gaussian. We will nevertheless assume that conditionally to $\mathbf{I}_{0:k}$, $\mathbf{w}_k$ is a white noise of zero mean and covariance Q_k.

Estimating the motion parameters on the whole image grid brings a global information on the point motion. Such type of information is crucial in case of lack of local information like in case of noise or occlusions. To determine if a point to be tracked belongs to the dominant motion model, a segmentation map of outliers associated to a global motion model [10] is used at the initial time.

3.2. Conditional measure equation

In order to avoid any drif situations, the conditional measure equation is built to obtain a goodness of fit criteron between a reference template and the current image.

At time k, we assume that $\mathbf{x}_k$ is observable through a matching process whose goal is here to provide in image $\mathbf{I}_k$ the point which is the most similar to the initial point $\mathbf{x}_0$ in a reference template $\widetilde{\mathbf{I}}_0$. The result of this process corresponds to a correlation peak and defines the measure $\mathbf{z}_k$ of our system. The reference template is defined as the initial image $\mathbf{I}_0$, which has been eventualy updated by registration in case of large geometric and/or photometric deformations around the tracked feature. The template is updated when the uncertaincy on the estimate is small, i.e. the eigenvalues of $\Sigma_{k|k}$ are under a given threshold.

Several matching criteria can be used to quantify the similarity between the target point and the candidate point. The conservation of the intensity pattern assumption has simply brought to consider the sum-of-squared-differences

(SSD). The measure $\mathbf{z}_k$ is achieved such as:

$$\mathbf{z}_k = \arg\min_{\mathbf{z}} \underbrace{\sum_{\mathbf{y}\in\mathcal{W}} [\widetilde{\mathbf{I}}_0(\mathbf{x}_0+\mathbf{y}) - \mathbf{I}_k(\mathbf{z}+\mathbf{y})]^2}_{r_k(\mathbf{z})}. \quad (3)$$

$r_k(\mathbf{z})$ is the residual associated to point $\mathbf{z}$. We suppose that this measure carries enough information about the state of the point being tracked to be able to write that $\mathbf{x}_k = \mathbf{z}_k$ apart from a white Gaussian noise $\mathbf{v}_k$. The SSD surface is in fact modeled by a 2D-Gaussian distribution, whose covariance accounts for the measure noise covariance. This covariance accounts for a confidence measure on the matching.

As the system is composed of a linear measure equation and a linear state equation conditionally to the image sequence, it can be solved with the proposed CLF.

3.3. Automatic measure noise covariance

A good estimation of the measure noise covariance R_k is essential to make the tracker robust to corrupted observations. Indeed, many factors can affect the quality of the measure: large affine geometric or photometric deformations, occlusions, etc. To that end, we define an SSD surface, around the measure $\mathbf{z}_k$. To evaluate a confidence on the SSD result, our approach uses the idea of Singh and Allen [11]. Each value on the SSD surface, corresponding to an error distribution is transformed into a response distribution:

$$\mathcal{D}_k(\mathbf{z}) = \exp(-c\, r_k(\mathbf{z})), \quad (4)$$

where c is a normalisation factor, such as $\sum_{\mathbf{z}\in\mathcal{W}'} \mathcal{D}_k(\mathbf{z}) = 1$. We assume that this distribution corresponds to a probability distribution of the true match location. The covariance matrix R_k associated to the measure $\mathbf{z}_k$ is constructed from the distribution (4):

$$R_k = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{xy} & \sigma_{yy} \end{pmatrix}, \quad (5)$$

with $\sigma_{uv} = \sum_{\mathbf{z}\in\mathcal{W}(\mathbf{z}_k)} \mathcal{D}_k(\mathbf{z})(u-u_k)(v-v_k)$, $\mathbf{z}_k = (x_k, y_k)^t$ and $\mathbf{z} = (x, y)^t$ belonging to the support around $\mathbf{z}_k$. Such a modelization allows to define an adaptative ellipse of uncertainty of the match location $\mathbf{z}_k$.

Nevertheless, a problematic issue occurs when the response distribution can not be approximated by a Gaussian distribution. This is the case when the SSD surface exhibits numerous significant peaks. The covariance construction described in (5) is then not relevant. This may happen in case of occlusions and particularly for highly textured areas. The corresponding $\mathcal{D}_k$ surface may then be better fitted by a uniform distribution than a Gausian distribution. To overcome a mis-approximation, a Chi-square "goodness of fit" test is realised, in order to check if the response distribution is better approximated by a normal or a uniform law. In this latter case, the diagonal terms of R_k are fixed to infinity, and the off-diagonal terms are set to 0. This test improves significantly the results, by allowing a robust detection of occlusions and other ambigous situations.

4. RESULTS

In this section, we present some experimental results on real-world sequences to demonstrate the efficiency of the proposed point tracker. We compare it to the Shi-Tomasi-Kanade (STK) tracker and to a robust differential method (RDM), which corresponds to our dynamic described § 3.1. We also compare it to a CONDENSATION-like nonlinear filter (NLF) [5], with a similar system to the one used for the CLF.

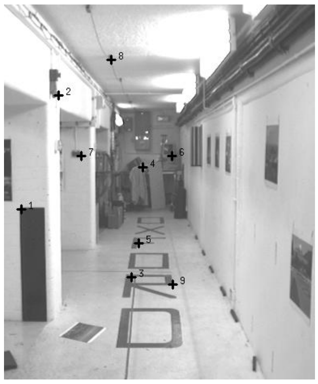
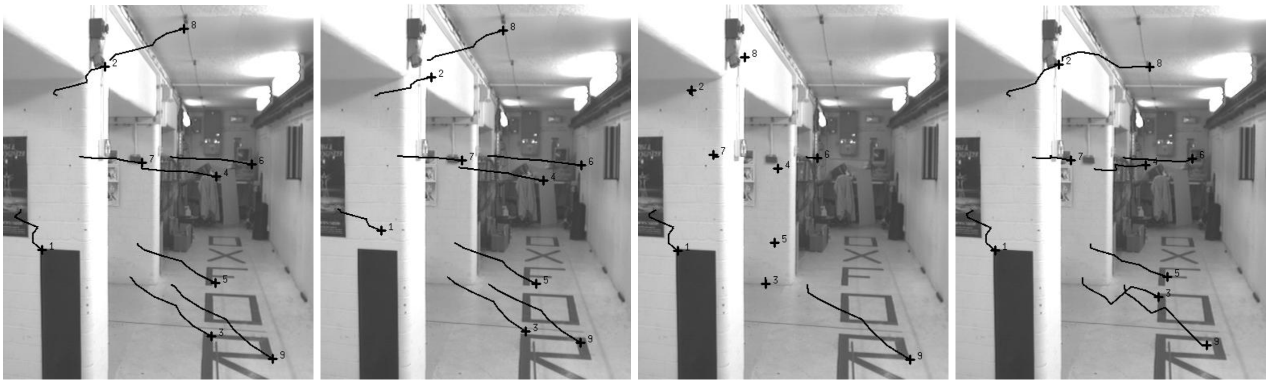


Fig. 1. *Sequence Corridor, initial points on image # 0*

The first sequence, **Corridor**, is a 7-frame (512 × 512 pixels) sequence, which constitutes, due to depth discontinuities and large motions, an extreme case for a global affine motion model. The initial points are presented figure 1. The complete trajectories provided by CLF, STK,NLF and RDM are presented figure 2. In such a sequence, it can be noticed that the STK leads to good tracking results only for two points and loses the others on frame 2. On the opposite, for the CLF, the trajectories of all the points are well-recovered. On this figure, we can also observe that the NLF gives less good results that the CLF, for a higher computational cost. We believe that in the one hand, the global dynamic on which we rely on gives a quite good prediction for the different points (even if it is a crude model in that case), and on the other hand, the matching measures authorize to correct the deficiency of such a motion model. This can be checked by looking at the results of the tracker built from the dynamic (RDM): On this result, it can be observed that a dominant motion model constitutes a quite rough motion model for some points (see points 1,2,9). Another illustration of these comments is presented on figure 4(a) which shows the comparative trajectories of the tested



CLF (# 6) RDM (# 6) STK (# 6) NLF (# 6)

Fig. 2. *Sequence Corridor*

algorithms, and a ground truth given by a user of points 2. As it can be observed, there is a significant deviation between the ground truth and the RDM trajectory. The poor result of the STK can also be observed on the same graphic.

A second result of the CLF is presented on **Hangars**, a 10-frame (512 $\times$ 512 pixels) noisy military sequence presenting a global chaotic motion. See for instance representation of point 13 trajectories on figure 4(c). A comparison between CLF, STK and NLF is presented figure 3. In such a sequence we can remark that STK leads to poor tracking results. This is particularly true for points which may not be easily identified by characteristic luminance patterns (corner points etc.). This is indeed a well-known deficiency of such a tracker. On the opposite, for the CLF, the trajectories of all the points are well-recovered. This is even more noteworthy that the sequence is noisy and the motion complex (see figures 4(b,c) for a representation of point 5 and point 13 trajectories). It is clear that in such a context, having a global information on top of local information is crucial. As for the sequence Corridor, in these figures, we can also observe that the NLF gives less good results than the CLF, for a higher computational-time. The last sequence, **Garden**, is used to demonstrate the robustness to occlusions. This is a sequence of 27 frames (240 $\times$ 360) showing a garden and a house occluded by a tree. Apart from the tree, this sequence presents a global translational motion. Figure 5 shows the CLF results for the background points. Following the trajectories of points moving behind the tree, we can remark that, without specifying any occlusions scheme, the tracker recovers the point locations after they have been hidden. Indeed, the adaptive covariance noise, estimated from the sequence allows the conditional trackers to be resistant to occlusions.

5. CONCLUSION

In this paper, we propose a Conditional Linear Filter for point tracking in image sequence. This tracker has the particularity to deal with *a priori*-free systems, which entirely depend on the image data. The particular system considered for point tracking application uses both global and local pieces of information on the point motion, by combining a dynamic relying on a differencial method and measures based on a correlation criterion. The resulting tracker has led to very good tracking results for trajectories undergoing abrupt changes in noisy situations. This algorithm has been shown to be resistant to occlusions.

6. APPENDIX: EXPRESSION OF $E^*_W[X|Y]$

Reminding that, for two arbitrary random vectors Y and W:

$$\begin{aligned} E[\|Y\|^2|W] &= E[Y^tY|W] = E[tr\{YY^t\}|W] \\ &= tr\{cov(Y,Y|W)\} + E[Y|W]^t E[Y|W], \end{aligned}$$

where tr means the trace of the matrix in braces, and denoting for arbitrary random vectors X, Y and W

$$\begin{aligned} \Sigma_{X,Z|W} &\triangleq cov(X,Z|W) \\ &= E[(X - E[X|W])\,(Z - E[Z|W])^t|W] \\ &= E[XZ^t|W] - E[X|W]\,E[Z|W]^t, \end{aligned}$$

after few manipulations, one can write:

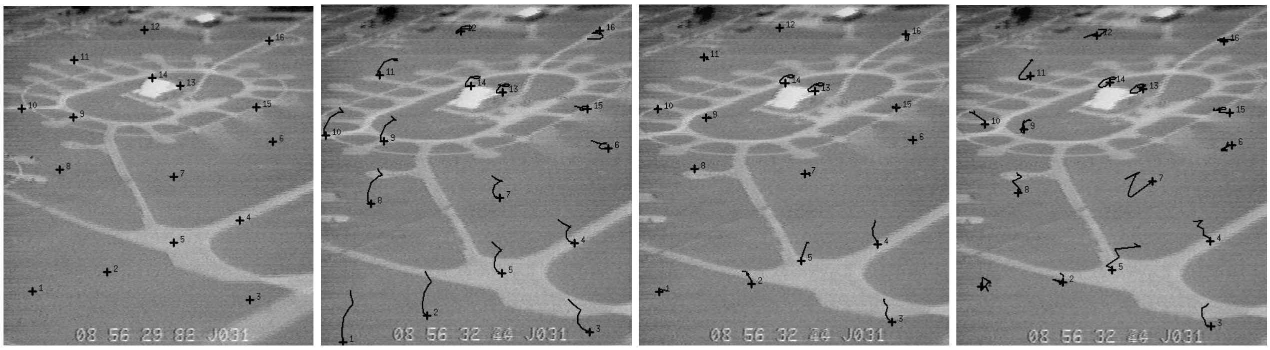
$$\begin{aligned} & E[\|X - A\,Z - B\|^2|W] \\ & = tr\{(A - \Sigma_{X,Z|W}\Sigma^{-1}_{Z,Z|W})\Sigma_{Z,Z|W} \\ & \qquad (A - \Sigma_{X,Z|W}\Sigma^{-1}_{Z,Z|W})^t\} \\ & + \|E[X|W] - AE[Z|W] - B\|^2 \\ & + tr\{\Sigma_{X,X|W} - \Sigma_{X,Z|W}\Sigma^{-1}_{Z,Z|W}\Sigma_{Z,X|W}\}. \end{aligned}$$

All the three terms are nonnegative.
$E[\|X - A\,Z - B\|^2|W]$ reaches its minimum for:

$$A = \Sigma_{X,Z|W}\,\Sigma^{-1}_{Z,Z|W},$$

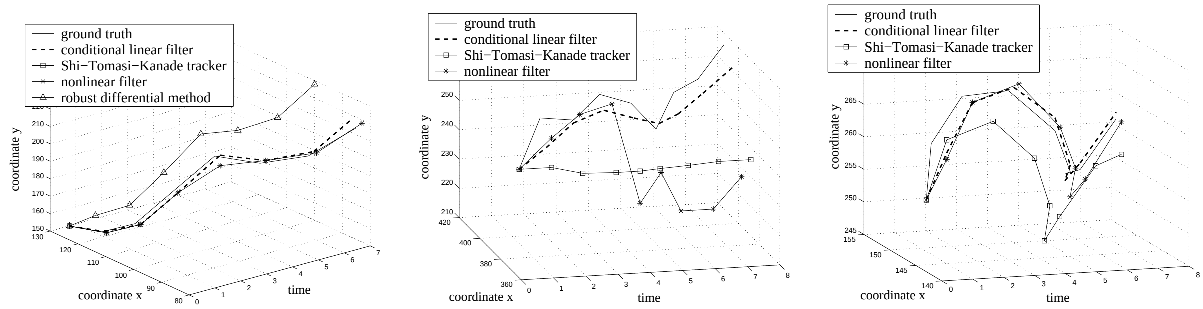
$$B = E[X|W] - \Sigma_{X,Z|W}\,\Sigma^{-1}_{Z,Z|W}\,E[Z|W].$$

We deduce the expression of $E^*_W[X|Z]$.



initial points on image # 0 CLF (# 9) STK (# 9) NLF (# 9)

Fig. 3. *Sequence Hangars*



(a) Sequence Corridor, point 2 (b) Sequence Hangars, point 5 (c) Sequence Hangars, point 13

Fig. 4. *Sequence Corridor and Hangars, comparison of trajectories*

initial points on image # 0 # 7 # 20 # 23

Fig. 5. *Sequence Garden*

7. REFERENCES

[1] P. Aschwanden and W. Guggenbühl, "Experimental results from a comparative study on correlation-type registration algorithms," *In W. Förstner and St. Ruwiedel, editors, Robust Computer Vision*, pp. 268–289, 1992.

[2] J. Shi and C. Tomasi, "Good features to track," in *CVPR*, 1994, pp. 593–600.

[3] N.P. Papanikolopoulos, P.K. Khosla, and T. Kanade, "Visual tracking of a moving target by a camera mounted on a robot : a combination of control and vision," *IEEE Trans. on Robotics and Automation*, vol. 9:1, pp. 14–35, 1993.

[4] F. Meyer and P. Bouthemy, "Region-based tracking using affine motion models in long image sequences," *CVGIP:IU*, vol. 60:2, pp. 119–140, 1994.

[5] M. Isard and A. Blake, "Condensation – conditional density propagation for visual tracking," *IJCV*, 1998.

[6] M. Isard and J. MacCormick, "Bramble : A bayesian multiple-blob tracker," in *ICCV*, 2001, vol. 2, pp. 34–41.

[7] P. Pérez, C. Hue, J. Vermaak, and M. Gangnet, "Color-based probabilistic tracking," in *ECCV*, June 2002, pp. 661–675.

[8] Y. Bar-Shalom and T.E. Fortmann, *Tracking and Data Association*, Accademic Press, 1988.

[9] J.M. Odobez and P. Bouthemy, "Robust multiresolution estimation of parametric motion models," *J. of Visual Comm. and Image Representation*, vol. 6:4, pp. 348–365, 1995.

[10] M. Irani, B. Rousso, and S. Peleg, "Computing occluding and transparent motions," *IJCV*, vol. 12:1, pp. 5–16, 1994.

[11] A. Singh and P. Allen, "Image-flow computation : An estimation-theoretic framework and a unified perspective," *CVGIP: IU*, vol. 56:2, pp. 152–177, 1992.